マルチメディア通信と分散処理　106－17
コンピュータセキュリティ　16－17
（２００２．２．１４）

DNA バイオメトリックス本人認証システム

板倉　征男*‡　　長嶋　登志夫‡　　辻井　重男*

*中央大学　研究開発機構
‡ NTT データテクノロジ(株)

筆者らは個人識別 ID のために用いる DNA 情報として，全塩基配列のなかで STR（Short Tandem Repert）と呼ばれる座位を複数箇所指定しそこで得られる繰返し回数情報を一定の順序で並べて個人識別子（以下 DNA 個人 ID と呼ぶ）を生成する事を提案し，実用化のための数々の基本的考察を行った．本論文ではこの DNA 個人 ID を生体識別子としたバイオメトリックス本人認証及びバイオメトリックス署名方式について実現するシステムを提案する．また，提案の方式を検証するために，500 人以上の提供者の協力を得て実際の人体の DNA を採取し，本方式によりバイオメトリックス本人認証が可能であることを検証した．実用化のために，リアルタイムによる DNA 分析装置の開発が条件となるが，本装置の実現までの間は 2 枚の IC カードを用いて認証を行う方式を代替方式として提案する．

Biometric Personal Authentication Using DNA Data

Yukio Itakura*‡　　Toshio Nagashima‡　　Shigeo Tsuji*

* Chuo University, Research and Development Initiative
‡ NTT Data Technology Corporation

This paper focuses on DNA data that can produce unique digital information for the purpose of personal identification. It discusses how to collect that information and describes the procedures for processing it to generate identification (ID) data. Based on statistical theory, this paper demonstrates that such information can be applied adequately to personal identification. In addition, the paper proposes a biometric personal verification/identification and digital signature system, and describes its implimertation overall features. In this paper we build a public key encryption method that incorporates DNA data into a secret key and authenticates individuals according to the public key encryption scheme.

１．はじめに

生体情報の中で DNA 情報はその採取・分析が難しく，プライバシの問題もあるので，これまでバイオメトリックス認証の要素としては鑑定などのごく限られた用途に限られていた．

しかし DNA 情報は，原理的にデジタル情報であり個人差の著しい部分，例えば STR（Short Tandem Repeat）と呼ばれる数個の塩基配列の繰返し回数の採取箇所を多重化すれば容易に識別精度を上げることが出来るので，これを個人識別子（DNA 個人 ID と呼ぶ）として応用すればこれまでの指紋や虹彩等によるものにはない新たな用途が考えられる．

プライバシの問題も ID のマッピングを行い，暗号鍵に組込むことによりブレークスルーを図ることが出来る．

本論文では筆者らが提案する DNA 個人 ID を用いた DNA バイオメトリックス本人認証及び署名方式について提案し，実証実験を行った結果を報告する．

２．DNA 個人識別子の生成

２．１　DNA 情報の特徴と応用の動機

図１はバイオメトリックス認証から見た DNA 情報の特徴を従来の指紋などと比較したものである．

属性 ＼ 生体情報	DNA 情報	従来方式
識別子の情報元（属性，データ長）	DNA の STR 繰返し数の個人差（デジタル情報，20 バイト）	指紋，虹彩，網膜，音声等のパターンの個人差（アナログ情報，250～1500 バイト）
識別アルゴリズム	数値（DNA 個人 ID）比較	パターンマッチング
識別精度	$\sim 10^{40}$（STR30 多重のときの実用的同値識別率）	$\sim 10^{-6}$（指紋の場合の実用的他人受容率）
識別時間	3H～（現状ではリアルタイム処理が不可）	～数秒（オンサイトによるリアルタイム処理可）

図１　DNA 情報を用いたバイオメトリックスの特徴

バイオメトリックス本人認証にDNA情報を応用する動機は,それが次の3つの属性を有することにある.

(1) 情報元の精度と絶対性

DNA情報から提案する方法で個人識別子（DNA個人ID）を生成するとIDが同値となる確率は，STRの多重度を上げれば指数的に減少する.n=30とすると，あらゆる組合せに対する実用的同値識別率は10^{-10}程度となり高精度の識別が可能である.[1]

しかも後述のような工夫をすればDNA個人IDは本人と1:1の確定数値となるので，様々なバイオメトリックス応用システムを考えることが出来る.

(2) 生体情報特徴量としてのコンパクト性

後述するDNA個人IDは従来のバイオメトリックスシステムにおける生体識別用特徴量情報に相当する.従来250バイト（指紋の例）～1,500バイト（声紋の例）を要した特徴量情報に対してDNA個人IDは20バイトである．これはDNA個人IDの元がデジタル情報であることの大きな利点である.

(3) 情報元の経年不変性と安定性

DNA情報は人間の細胞全てが同じ塩基配列，つまり同じ情報であり，一生不変とされている．また塩基そのものは4種類の無機質の化合物であるが，これらは何十年前の骨からDNAが採取されるようにきわめて安定な物質である．これらは他のバイオメトリックスの弱点をカバーする優れた属性を有していると言える.

2．2　DNA個人IDの生成方法

DNA個人ID α_Aは各STRローカス（座位）のL_i（$j \| k, j \leqq k$）を順に並べた配列で次のように生成する.

$$\alpha_A = L_1 \| L_2 \| L_3 \| \cdots \| L_n \qquad (1)$$

但し，L_iはローカスi番目のSTR繰返し回数の（$j \| k$）$_i$を示す.

生成されたα_Aは,後述するように一定の確率で一意性のある個人識別情報となる.[1]

STRローカスの多重度をn，生成したDNA個人IDを使用する集団の人数をNとすると，N人のあらゆる組合せに対する同値確率P（これを実用的同値識別率と呼ぶこととする）は,

$$P \quad \frac{1}{2}N(N-1)P_n \qquad (2)$$

となる．ここでP_nは実証実験等から$P_n \quad 10^{-n}$が得られている.[1]

N人の集団で提案するDNA個人IDを使って情報セキュリティシステムを設計する場合，実用的同値識別率Pの値を$\approx \frac{1}{N}$と設定すると，必要なSTR多重度nは上記の実験式で求められる．（n，N，P）の関係について数例を示すと，図2のようになる．DNA個人IDは，実用的同値識別率以下の確率で発生する同値のDNA個人IDの間や一卵性双生児の間で同値となる，もしくは同値となる可能性がある.

このような場合においても識別を可能とするために，何らかの対策が必要である．このため本方式では登録しようとするDNA個人IDを登録済の他人の数値と比較し，同値となった場合は3．1節の(4)式生成の過程で乱数を振り直して別の値になるような処置を行うこととしている.

n (STR multiplicity)	N (Number of people in the group)	P (Practical matching value recognition rate)
15	10^5	$\frac{1}{2}10^{-5}$
21	10^7	$\frac{1}{2}10^{-7}$
24	10^8	$\frac{1}{2}10^{-8}$
30	10^{13}	$\frac{1}{2}10^{-10}$

図2　STR多重度nと実用的同値識別率Pの関係

2．3　DNA個人IDのマッピング

生成されたDNA個人ID α_Aはハッシュ関数処理を行う．その出力をδ_Aとする.

$$\delta_A = h(\alpha_A) \qquad (3)$$

ここでh( )としては汎用一方向性ハッシュ関数であるSHA(Secure Hash Algorithm)-1を適用する.[2] α_Aはn(STRの多重度)=15の場合10^{50}程度,n=30の場合10^{100}程度のビット長の情報となる.SHA-1の場合入力が2^{64}ビット長未満のビット列で出力が160ビット長のメッセージダイジェストとなる一方向性関数演算を行うので，本方式に採用できるハッシュアルゴリズム関数である．なおこのh( )は公開する．公開することによりδ_Aの普遍性が保証される．即ち，同一人物のδ_AはどこでDNAを採取しても同じ値となる．ハッシュ関数は一方向性関数であるから，個人のプライバシであるα_A(DNA個人ID)を秘匿する意義がある.

3．　DNA個人IDの暗号鍵への組込み

3．1　暗号鍵へ組込む意義

DNAバイオメトリックス本人認証方式及び後章のバイオメトリックス署名方式をシステム化するに当たり，基盤となる個人情報であるDNA個人IDについて,

次の理由でこれをまず暗号鍵に組込む.

(1) プライバシの保護

DNA 個人 ID は，人間の基本的生体情報であり，プライバシ保護の考慮を第一義としてシステム化を論じなければならない.

従って生成した DNA 個人 ID は例えハッシュ関数を通したものでも識別判定の情報としてそのまま適用することは，倫理上好ましくないこと．このため，DNA 個人 ID 情報が組込んであることを証明することは出来るが，DNA 個人 ID 情報そのものは秘密とすることが出来るようなしくみが必要である．本論文ではδ_A を秘密鍵に組込み，そのペアとなる公開鍵を生成して，それを CA（認証局）に登録する方式を提案する.

バイオメトリックス本人認証は，DNA 個人 ID を生体情報で判別するのではなく，ここで提案する生体情報組込み型公開鍵が生成できるか否かで判別する方式で行う.

(2) 直接組込可能な DNA 情報の属性

従来のバイオメトリックスの特徴量情報は，ID としては相対的な情報であり，かつデータ長も大きいので鍵に直接組込むことは困難であった．一方これに相当する DNA 個人 ID は個人識別のための絶対的な情報であり，かつデータ長も 8 バイトに圧縮出来る属性を有するため，暗号鍵に直接組込むことが実現可能となる.

(3) 生体情報 DB の割愛

暗号鍵に組込むことにより，認証の為のリアルタイムの生体情報 DB を特別に構築する必要はなくなり，既存の PKI のしくみの中でバイオメトリックス本人認証を実現することが可能となる.

3．2　秘密鍵への組込みと公開鍵の生成方法[3)]

(1)秘密鍵への組込み

ここでは公開鍵暗号方式における秘密鍵に DNA 個人 ID を組込むことを述べる.

秘密鍵を X_A とする. X_A は 160 ビット長程度のビット列を考える．X_A は次の計算により生成する.

秘密鍵：X_A（2^{160} ビット長程度のビット列）

$$X_A = \delta_A \oplus r_A \qquad (4)$$

ここで r_A は 160 ビット長程度のビット列の個人秘密乱数であり，登録用専用端末で生成する.

r_A を必要とする理由は，生体情報としてのα_A は，他人の DNA を盗むことが可能なので，そのままでは秘密情報とはなり得ないこと，従って個人の秘密情報 r_A を加えて秘密鍵とする必要があるためである.

(2)公開鍵への組込み

ここでは離散対数問題に基づく公開鍵暗号である ElGamal 暗号を採ることとし，前項で生成した秘密鍵から次のようにして公開鍵 Y_A を生成する.

公開鍵：Y_A

$$Y_A = g^{X_A + r_A} \pmod p \qquad (5)$$

（p は大きな素数，g は位数 p の原始元であり，システムに共用な値である）

認証局(CA)への登録：

Y_A，g^{r_A} 及び個人情報を登録する.

生成した公開鍵は，他の関係情報と合せて認証局(CA)に登録する．この際 CA は過去に登録済の CA を調べ，同一の公開鍵がある場合は r_A を生成し直して新たな Y_A を生成し登録し直すことを要求する．これにより同一の DNA 個人 ID を有する者が現れた場合の課題が解決出来る.

4．DNA バイオメトリックス認証方式の原理

4．1　DNA バイオメトリックス本人認証方式

本人の個人情報と前項で述べた秘密鍵及びそのペアとなる公開鍵などを耐タンパー性のある IC カードに記録し，携行する．これをバイオメトリックス実印カードと名づける.

ここでは図３により，この IC カードの持主の正当性を検証するための本人認証方式について説明する.

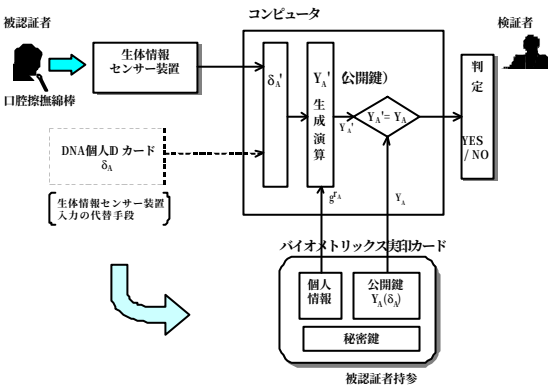


図３　DNA バイオメトリックス本人認証方式の原理

本人認証を行うシステムは，生体情報（DNA）センサー装置を備え，本人の口腔を軽く擦った綿棒を入力情報としてカードの持ち主が本人であることをバイオメトリックス認証技術により確認する．即ち綿棒に付着した口腔細胞の分析により DNA 個人 ID を生成し，これから本人の公開鍵が生成できるか否かをテストする.

今自分は A であると称する人物 Aから採取したα_A'

にハッシュをかけた DNA 個人 ID をδ_A' とする．これは生体情報センサー端末からの生体情報入力となる．

一方バイオメトリックス実印カードから読み出すg^{rA}，及びシステム共通情報である g，p を使い，上記δ_A'から自分は A であると称する人物の公開鍵 Y_A'を生成する．計算方法は以下の通りである．

$$Y_A' = g^{X_A'} = g^{\delta_A' + rA} = g^{\delta_A'} \cdot g^{rA} \pmod p \qquad (6)$$

生成した Y_A'がバイオメトリックス実印カードに記録されている Y_A と等しいか検証する．

$$Y_A = Y_A' ?$$

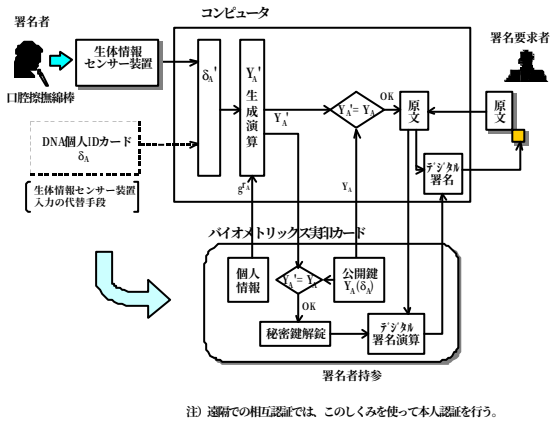
一致すれば A'は A 本人であることが認証される．もし A 以外の人物が偽って A を主張しても，自分のδ_A'で Y_A が生成できなければ本人になりすますことは不可能である．

現状技術では生体情報センサー装置は分析結果を出すのに 3 時間以上かかり，リアルタイムによるオンサイトチェックには使えない．このため自分の α_A から生成した$\delta_A = h(\alpha_A)$を別の IC カード（DNA 個人 ID カードと名づける）に記録しておき，この 2 枚目のカードで生体情報センサー装置からの生体情報入力を代行させる方式が考えられる．

以上は IC カードの持主の正当性を確認する方式である．これに対してネットワークを介した相互認証により本人確認を行う場合，即ち遠隔での本人認証については次節で述べる DNA バイオメトリックス署名方式を使い，検証者からのチャレンジ電文にデジタル署名をつけてリターンする交信プロトコルにより実現する．

4．2　DNA バイオメトリックス署名方式

図４により DNA バイオメトリックス署名方式を説明する．



注）遠隔での相互認証では、このしくみを使って本人認証を行う。

図４　DNA バイオメトリックス・デジタル署名方式の原理

まず「バイオメトリックス実印カード」の持ち主の正当性を前節と同じ手順で確認し，OK となるとカードに記録された秘密鍵が解錠され，「バイオメトリックス実印カード」にあるデジタル署名プログラムが入力された文書情報に署名をつける．署名が終わると秘密鍵の施錠を行う．IC カードの中にデジタル署名機能を持つので，秘密鍵の情報を外に漏らさずに署名が出来る．

DNA バイオメトリックス署名では，自分の DNA 情報がデジタル署名に埋め込まれることになるので，血判を押印したような心理的効果も期待される．

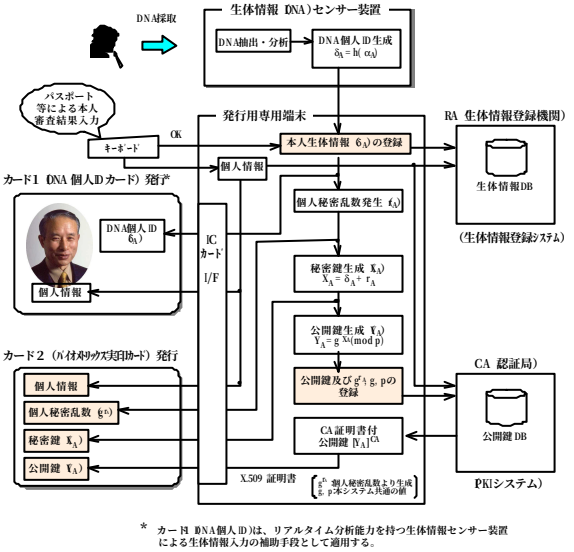
5．DNA バイオメトリックス・システムの具体的構成(4),(5)

5．1　DNA バイオメトリックス登録システム

本節では４章の方式を使って具体的に DNA バイオメトリックス・システムを構築する場合の具体例を提案する．

まず DNA の生体情報を登録するしくみが必要であるが，本方式では生体情報をそのまま扱うのではなく，暗号鍵に組込む方法を採るため，図５のような公開鍵暗号方式における秘密鍵と公開鍵の生成及び CA（認証局）に公開鍵を登録するのと類似なしくみとなる．この装置を発行用専用端末と呼ぶこととする．

CA には公開鍵 Y_A の他に暗号化された個人秘密乱数 g^{rA} を合せて登録する．この値はバイオメトリックス本人認証において，DNA 個人 ID δ_A から公開鍵 Y_A が生成出来るか検証演算を行うとき必要なものである．



＊ カード1（DNA個人ID）は、リアルタイム分析能力を持つ生体情報センサー装置による生体情報入力の補助手段として適用する。

図５　DNA バイオメトリックス登録システム

発行用専用端末で登録用に生成したδ_A，X_A，Y_A，g^{rA} 及びキーボードから入力した氏名，生年月日などの

個人情報は，バイオメトリックス実印カード（図中のカード２）に書き出し記録する．この IC カードは耐タンパー性のある構造であることが必要である．

バイオメトリックス実印カードは，これらの情報の他に Y_A の生成・比較演算及びデジタル署名演算機能を有するプログラムを内蔵する．前者はデジタル署名時に秘密鍵を解錠するために，また後者はデジタル署名時に秘密鍵を外に漏らさないために，IC カード内に内蔵するプログラムである．

生体情報センサー装置の代替または補完機能として，DNA 個人 ID カード（図中のカード１）を併用する場合は，これに δ_A 及び個人情報を書き出し記憶する．

５．２　DNA バイオメトリックス認証システム

DNA 情報を用いたバイオメトリックス本人認証及びバイオメトリックス署名を行うシステムの構成例を図６に示す．

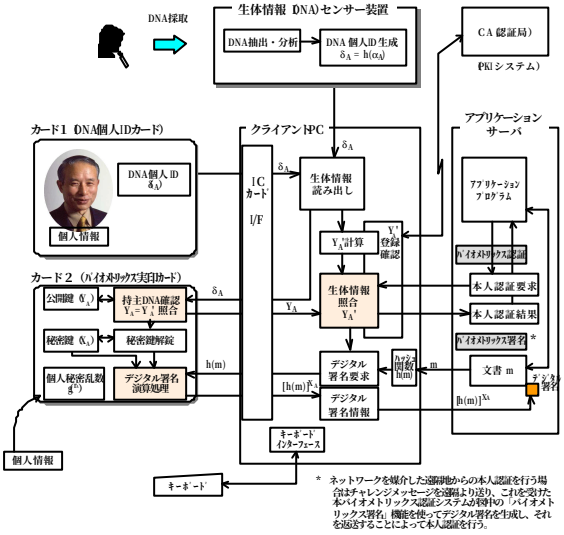


図６　DNA バイオメトリックス認証システム

ここではクライアント PC にこれらの２つの処理機能を持たせ，アプリケーションサーバにあるアプリケーションプログラムからこれらの機能をクライアント PC に要求し，処理を行わせる設計としている．

６．DNA バイオメトリックス・システムの実証実験

（１）実験システムの構築

本方式を検証するために，実験システムを構築した．但し生体情報センサー装置は，現在の技術ではリアルタイム処理が不可能であるために仮想のものとし，実際の生体情報（DNA）の分析は既存の法医学鑑定に用いる分析機械を用いてオフラインで行った．(1)

また実験システムではバイオメトリックス実印カード（カード２）の機能のうち公開鍵生成処理とデジタル署名処理のプログラムを除く機能を１枚の IC カードにインプリメントして実証実験を行った．

実験の第一ステップとして，本人認証機能のテストを行うこととし，1,000 人分の PKI（公開鍵基盤）機能を有する CA を実現した．

（２）実証実験の方法

本研究では，中央大学研究開発機構と東北大学大学院医学系研究科との共同研究協定を結ぶことにより，515 人の DNA 提供者の協力を得て DNA 個人 ID の統計的検証とバイオメトリックス本人認証システムの実用化の可能性について実証実験を行なった．

実証実験によって実 DNA から得られたデータは次の様な手順で解析を行い，検証を行った．

採取し生成した 515 人の DNA 個人 ID について同値確率を調べた．ローカス 15 多重をもとに生成した DNA 個人 ID の同値のものの存在は，全ての組合せにおいては 0 となった．

次にローカスの多重度を 15から順次減らしていき，それに応じた DNA 個人 IDの同値率を調べて行くと，ローカスの多重度が 5 の時はじめて同値のものが表われ，以小の多重度のものについては同値の DNA 個人 ID が存在することが実証された．図７に採取した DNA による実証実験結果を示す．

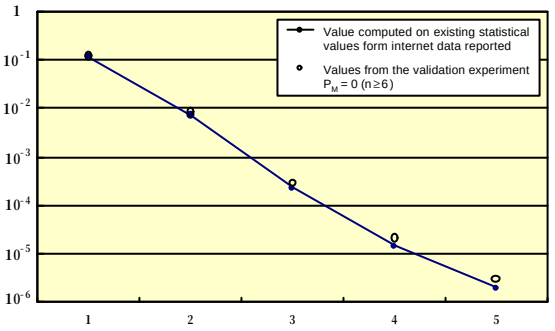


図７　採取した DNA による実証実験結果
（STR ローカス多重度と DNA 個人 ID の同値確率の関係）

実証実験に基づく同値確率の STR 多重度とインターネット公開データ(6)に基づくそれとはよく近似しており，実験式として

$$Pn \quad 10^{n} \qquad (7)$$

を得ることが出来た．これにより DNA バイオメトリックス本人認証は式(2)のような実用的同値識別率の精度で判別可能なことが検証出来た．

7．考察

7．1　分析時間の課題

DNA 個人 ID を適用してバイオメトリックス認証を行う場合，生体情報採取後の実用的な分析時間は数秒のオーダによる処理が要求される．現状の法医学分野で使用される分析機材では，最新のものでも 3 時間以上を要しているので，情報セキュリティシステムで実用化するには，革新的な技術によるブレークスルーが必要である．

上記のようなオンサイト・リアルタイム分析が不可能な現状では，要求仕様を満たす端末が開発されるまでの代替手段として，δ_A（DNA 個人 IDα_A にハッシュをかけた情報）を専用の IC カードに記憶し，これを使って本人の生体情報を入力する方法を５．１節で提案した．

法廷等でバイオメトリックス（DNA）による厳密な本人認証を行う場合は，時間をかけて生体情報を実際に分析するという担保を有することがポイントである．

7．2　倫理的課題

本提案による DNA 個人 ID は，DNA 情報のうち人体の構造や病因に関与しない，いわゆる遺伝子領域以外の部分でマイクロサテライトといわれる STR 情報を用いるので，個人の秘密情報には関与するものではない．しかしながら指紋のように個人識別が可能な本人固有な情報を取扱うので，プライバシ保護について十分考慮する必要がある．本提案ではプライバシ対策として次の２つを考えた．

その１つは DNA 個人 ID にハッシュ関数処理を行い，一方性のマッピングを行った情報を識別子（δ_A）として使用すること，その２つは DNA 個人 ID を秘密鍵及びそのペアとなる公開鍵に組込んでしまい，以降は公開鍵暗号方式の機能を使って署名や認証を行う．

暗号鍵に組込まれた生体情報が自分のものか否かは，その情報を直接相手に示さなくても４．１節に述べた方法で証明することが出来る．

また総合的見地から視ると，指紋と同様生体情報を個人番号と同等に扱い，そのような個人情報で人間を管理することの是非について討議する必要がある．

テロリズムが人間の安全な生活を脅かしている昨今において，適切な倫理法の下で本方式が究極の個人認証システムとして正しい運用が行われれば，21 世紀の新しい情報セキュリティシステムとして威力を発揮することが期待出来る．

8．まとめ

本論文は DNA 個人 ID を応用して情報セキュリティシステムにおけるバイオメトリックス認証及びバイオメトリックス署名を実現するための基本的課題について検討した．

特にプライバシ保護の観点からハッシュ関数処理を行うことや，暗号鍵に組込むことを考察し，課題解決のための方法を提案した．また提案方式を検証するため 500 人以上の DNA 提供者の協力を得て実証実験を行った．その結果，提案する方式の実用性について基本的な検証を得ることが出来た．

最後に倫理的課題と社会システムとして討議すべき問題提起を行った．

謝辞

本研究の一部は，通信・放送機構の「超楕円暗号を核とした高性能セキュリティ機能の実現と電子社会システムへの応用プロジェクト」の支援により行いました．また本研究の実証実験の実施に当たり，東北大学法医学部門の舟山眞人教授及び橋谷田技官に多大な御協力をいただきました．合わせて深謝申し上げます．

参考文献

(1) 板倉征男，橋谷田真樹，長嶋登志夫，辻井重男：DNA-ID の統計的検証，信学技報，vol.101, No.214, ISEC2001-19, pp.1-7 (July 2001).
(2) 岡本龍明，山本博資：現代暗号，産業図書，pp.189-195（1997）．
(3) 辻井重男，板倉征男，山口浩，北沢敦，齋藤真也，笠原正雄：生体情報が秘密鍵に埋め込まれた構造を有する公開鍵暗号方式，電子情報通信学会シンポジウム，SCIS2000, D07 (Jan. 2000)．
(4) 日本自動認識システム協会編：これでわかったバイオメトリックス，オーム社，p.94-102，119-126（2001）．
(5) Isobe, Y., Seto, Y., Kataoka, M.：Development of Personal Authentication System Using Fignerprint with Digital Signature Technologies, IEEE Proceedings of the 34th Hawaii International Conference on System Science (2001).
(6) http://www.uni-duesseldolf.de/WWW/MedFak/Serology/database.html